

NTUTH-99-100
October 1999

A note on the solutions of the Ginsparg-Wilson relation

Ting-Wai Chiu

Department of Physics, National Taiwan University
Taipei, Taiwan 106, Republic of China.
E-mail : twchiu@phys.ntu.edu.tw

Abstract

The role of R in the solutions of the Ginsparg-Wilson relation is discussed.

PACS numbers: 11.15.Ha, 11.30.Rd, 11.30.Fs

In this note, we would like to clarify some seemingly subtle issues pertaining to the role of R in the solutions of the Ginsparg-Wilson relation [1]

$$D\gamma_5 + \gamma_5 D = 2D\gamma_5 R D . \quad (1)$$

Here R is a positive definite hermitian operator which is local in the position space and trivial in the Dirac space. Since we can sandwich (1) by a left multiplier \sqrt{R} and a right multiplier \sqrt{R} on both sides of (1), and define $D' = \sqrt{R}D\sqrt{R}$, then D' satisfies

$$D'\gamma_5 + \gamma_5 D' = 2D'\gamma_5 D' \quad (2)$$

which is in the same form of Eq. (1) with $R = \mathbb{1}$. This seems to suggest that one can set $R = \mathbb{1}$ in the GW relation (1) and completely ignore the R dependence in the general solution of the GW relation. However, as we will see, if R is set to be the identity operator from the beginning, then some of the salient features of the general solution may be easily overlooked.

First, let us review some basics of the GW relation. The general solution to the GW relation (1) can be written as [2, 3]

$$D = D_c(\mathbb{1} + RD_c)^{-1} = (\mathbb{1} + D_c R)^{-1} D_c \quad (3)$$

where D_c is any chirally symmetric Dirac operator, i.e.,

$$D_c\gamma_5 + \gamma_5 D_c = 0 . \quad (4)$$

In order to have D reproduce the continuum physics, D_c is required to satisfy the necessary physical constraints [3]. The general solution of D_c has been investigated in ref. [4]. Conversely, for any D satisfying the GW relation (1), there exists the chirally symmetric D_c

$$D_c = D(\mathbb{1} - RD)^{-1} = (\mathbb{1} - DR)^{-1} D . \quad (5)$$

We usually require that D also satisfies the hermiticity condition¹

$$D^\dagger = \gamma_5 D \gamma_5 . \quad (6)$$

Then D_c also satisfies the hermiticity condition $D_c^\dagger = \gamma_5 D_c \gamma_5$, since R is hermitian and commutes with γ_5 . The hermiticity condition together with the chiral symmetry of D_c implies that D_c is anti-hermitian. Thus there exists one to one correspondence between D_c and a unitary operator V such that

$$D_c = (\mathbb{1} + V)(\mathbb{1} - V)^{-1}, \quad V = (D_c - \mathbb{1})(D_c + \mathbb{1})^{-1}. \quad (7)$$

¹ This implies that $\det(D)$ is real and non-negative.

where V also satisfies the hermiticity condition $V^\dagger = \gamma_5 V \gamma_5$. Then the general solution (3) can be written as [2, 3]

$$D = (\mathbb{1} + V)[(\mathbb{1} - V) + R(\mathbb{1} + V)]^{-1} \quad (8)$$

$$= [(\mathbb{1} - V) + (\mathbb{1} + V)R]^{-1}(\mathbb{1} + V). \quad (9)$$

On the other hand, if one starts from Eq. (2), then its solution is

$$D' = \frac{1}{2}(\mathbb{1} + V) \quad (10)$$

which agrees with (8) with $R = \mathbb{1}$. Then using the relation $D' = \sqrt{R}D\sqrt{R}$, one obtains

$$D = \frac{1}{\sqrt{2R}}(\mathbb{1} + V)\frac{1}{\sqrt{2R}} \quad (11)$$

However, (11) is in contradiction with (8) since the R dependence in (11) can be factored out completely while that of (8) can not. So, one of them can not be true in general. If we take the limit $R \rightarrow 0$, then (8) gives that $D \rightarrow (\mathbb{1} + V)(\mathbb{1} - V)^{-1} = D_c$, but (11) implies that $D \rightarrow \infty$. Since D_c is well defined (without poles) in the trivial gauge sector, it follows that (11) can not be true in general. The fallacy in (11) is due to the assumption that D' is independent of R . From (8), one can derive the following formula

$$D = \frac{1}{\sqrt{2R}}(\mathbb{1} + V')\frac{1}{\sqrt{2R}} \quad (12)$$

where

$$V' = \left[(\mathbb{1} + V)\sqrt{R} + (\mathbb{1} - V)\frac{1}{\sqrt{R}} \right]^{-1} \left[(\mathbb{1} + V)\sqrt{R} - (\mathbb{1} - V)\frac{1}{\sqrt{R}} \right] \quad (13)$$

$$= \left[\sqrt{R}(\mathbb{1} + V) - \frac{1}{\sqrt{R}}(\mathbb{1} - V) \right] \left[\sqrt{R}(\mathbb{1} + V) + \frac{1}{\sqrt{R}}(\mathbb{1} - V) \right]^{-1} \quad (14)$$

which is unitary and depends on R . This shows that D' actually depends on R and equals to $\frac{1}{2}(\mathbb{1} + V')$ rather than (10). Therefore it is *erroneous* to write the general solution of the GW relation in the form of (11), in which V' is replaced by V . Consequently, (11) may mislead one to infer that R does not play any significant roles in the locality of D , in particular when R is proportional to the identity operator. However, R indeed plays a very important role in determining the locality of D . Let us consider $R = r\mathbb{1}$ with $r > 0$. When $r \rightarrow 0$, $D \rightarrow D_c$ which must be nonlocal if D_c is free of species doubling and has the correct behavior in the classical continuum limit [5]. As the value of r moves away from zero and goes towards a finite value, D may change from a non-local operator to a local operator. This has been demonstrated in ref.

[6, 7]. Therefore, the general solution (3) of the GW relation can be regarded as a topologically invariant transformation (i.e., $\text{index}(D) = \text{index}(D_c)$) which can transform a nonlocal D_c into a local D . Conversely, the transformation (5) can transform a local D into the non-local D_c . For a given D_c , the set of transformations, $\{ \mathcal{T}(R) : D = D_c(\mathbb{I} + RD_c)^{-1} \}$, form an abelian group with parameter space $\{ R \}$ [7]. In general, for any lattice Dirac operator D (not necessarily satisfying the GW relation), we can use the topologically invariant transformation $D' = D(\mathbb{I} + RD)^{-1}$ to manipulate its locality.

The next question is whether we can gain anything (e.g., improving the locality of D) by using another functional form of R rather than the simplest choice $R = r\mathbb{I}$. We investigate this question by numerical experiments. For simplicity, we consider the Neuberger-Dirac operator [8]

$$D_h = \mathbb{I} + V, \quad V = D_w(D_w^\dagger D_w)^{-1/2} \quad (15)$$

where D_w is the Wilson-Dirac fermion operator with negative mass -1

$$D_w = -1 + \frac{1}{2}[\gamma_\mu(\nabla_\mu^* + \nabla_\mu) - \nabla_\mu^* \nabla_\mu] \quad (16)$$

where ∇_μ and ∇_μ^* are the forward and backward difference operators defined in the following,

$$\begin{aligned} \nabla_\mu \psi(x) &= U_\mu(x) \psi(x + \hat{\mu}) - \psi(x) \\ \nabla_\mu^* \psi(x) &= \psi(x) - U_\mu^\dagger(x - \hat{\mu}) \psi(x - \hat{\mu}) . \end{aligned}$$

The Neuberger-Dirac operator D_h satisfies the GW relation (1) with $R = 1/2$. Then Eq. (5) gives

$$D_c = 2 \frac{\mathbb{I} + V}{\mathbb{I} - V} . \quad (17)$$

Substituting (17) into the general solution (3), we obtain

$$D = 2(\mathbb{I} + V)[(\mathbb{I} - V) + 2R(\mathbb{I} + V)]^{-1} . \quad (18)$$

For a fixed gauge background, we investigate the locality of $D(x, y)$ versus the functional form of $R(x, y)$. For simplicity, we consider D in a two dimensional $U(1)$ background gauge field with non-zero topological charge, and we use the same notations for the background gauge field as Eqs. (7)-(11) in ref. [9]. Although it is impossible for us to go through all different functional forms of $R(x, y)$, we can use the exponential function

$$R(x, y) = r \exp(-m|x - y|) \quad (19)$$

as a prototype to approximate other forms by varying the parameters r and m . In the limit $m \rightarrow 0$, $R(x, y)$ is nonlocal, while in the limit $m \rightarrow \infty$, $R \rightarrow r\mathbb{I}$

which is the most ultralocal. Hence, by varying the value of m from 0 to $\alpha \gg 1$, we can cover a wide range of $R(x, y)$ of very different behaviors.

One of the physical quantities which are sensitive to the locality of D is the anomaly function

$$\mathcal{A}_n(x) = \text{tr} [\gamma_5(RD)(x, x) + \gamma_5(DR)(x, x)] \quad (20)$$

which can serve as an indicator of the localness of D . Since the Neuberger-Dirac operator is *topologically proper* for smooth gauge backgrounds, the index of D in (18) is equal to the background topological charge Q ,

$$\text{index}(D) = n_- - n_+ = Q. \quad (21)$$

This implies that the sum of the anomaly function over all sites on a finite lattice must be equal to two times of the topological charge [7]

$$\sum_x \mathcal{A}_n(x) = 2(n_- - n_+) = 2Q = \begin{cases} \frac{1}{16\pi^2} \sum_x \epsilon_{\mu\nu\lambda\sigma} F_{\mu\nu}(x) F_{\lambda\sigma}(x), & d = 4; \\ \frac{1}{2\pi} \sum_x \epsilon_{\mu\nu} F_{\mu\nu}(x), & d = 2. \end{cases} \quad (22)$$

This is true for any R since the index of D is invariant under the transformation (3), i.e., $\text{index}(D) = \text{index}(D_c)$. First we consider the gauge configuration with constant field tensors. If D is local, then we can deduce that $\mathcal{A}_n(x)$ is constant for all x . From (22), it follows that

$$\mathcal{A}_n(x) = \rho(x) \equiv \begin{cases} \frac{1}{16\pi^2} \epsilon_{\mu\nu\lambda\sigma} F_{\mu\nu}(x) F_{\lambda\sigma}(x), & d = 4; \\ \frac{1}{2\pi} \epsilon_{\mu\nu} F_{\mu\nu}(x), & d = 2, \end{cases} \quad (23)$$

where $\rho(x)$ is the Chern-Pontryagin density in continuum. Note that Eq. (23) also implies that $\mathcal{A}_n(x)$ is independent of R if D is local. Next we introduce local fluctuations to the constant background gauge field, with the topological charge fixed. Then we expect that (23) remains valid provided that the locality of D is not destroyed by the roughness of the gauge field. Therefore, in general, by comparing the anomaly function $\mathcal{A}_n(x)$ at each site with the Chern-Pontryagin density $\rho(x)$, in a prescribed background gauge field, we can reveal whether D is local or not in this gauge background. This provides another scheme to examine the locality of D rather than checking how well $|D(x, y)|$ can be fitted by an exponentially decay function. We will use both methods in our investigations. We define the deviation of the anomaly function as

$$\delta_D \equiv \frac{1}{N_s} \sum_x \frac{|\mathcal{A}_n(x) - \rho(x)|}{|\rho(x)|} \quad (24)$$

where the summation runs over all sites, and N_s is the total number of sites.

r	m	δ_D	index(D)
0.5	0.5	1.366	1
0.5	1.0	0.5360	1
0.5	2.0	0.2086	1
0.5	5.0	7.209×10^{-3}	1
0.1	5.0	6.554×10^{-2}	1
0.5	$\gg 1$	3.848×10^{-4}	1
1.0	$\gg 1$	1.618×10^{-4}	1

Table 1: The deviation of the chiral anomaly function δ_D [Eq. (24)] versus $R(x, y)$ with parameters r and m defined in Eq. (19), on a 12×12 lattice, in a constant background gauge field with topological charge $Q = 1$.

In Table 1, we list the deviation of the chiral anomaly function, δ_D , versus $R(x, y)$ with parameters r and m defined in Eq. (19), on a 12×12 lattice, in a constant background gauge field with topological charge $Q = 1$. The last two rows with $m \gg 1$ corresponds to $R = r\mathbb{I}$, and they have the smallest deviations. They are both local, which can be checked explicitly by plotting $|D(x, y)|$ versus $|x - y|$, as shown in Fig. 1. For $r = 1/2$ (the row on top of the last row), it corresponds to the Neuberger-Dirac operator. In the first row, $R(x, y)$ is nonlocal since $m = 1/2 < 1$. It produces a nonlocal D , as shown in Fig. 2, so the resulting chiral anomaly is very different from the Chern-Pontryagin density and thus δ_D is very large. Now we increase m to 1, 2 and 5 successively, then R and D both become more and more local, as shown in Fig. 3, thus δ_D becomes smaller and smaller, as shown in the second, third and fourth rows of Table 1. These results indicate that *a nonlocal R does not produce a local D , and a local R does not make D more local than that using $R = r\mathbb{I}$* . For $m = 5.0$, if we decrease r to 0.1, then D becomes very nonlocal, as shown in Fig. 4. Consequently, its δ_D (in the fifth row) is about 10 times larger than that of $r = 0.5$ (in the fourth row). This suggests that *on a finite lattice, r cannot be too small, otherwise D will become nonlocal*.

Our numerical results listed in Table 1 as well as those plotted in Figs. 1-4 strongly suggest that *we do not gain anything by using other functional forms of $R(x, y)$ than the simplest choice $R(x, y) = r \delta_{x,y}$* . However, the value of r plays the important role in determining the localness of D . We have also tested other functional forms of $R(x, y)$ as well as many different gauge configurations. The results from all these studies are consistent with the conclusion that *the optimal choice for R is $R(x, y) = r \delta_{x,y}$* .

Now we come to the question concerning the range of proper values of r . We have already known that r cannot be zero or very small, otherwise D is nonlocal. On the other hand, r cannot be too large, otherwise D is highly peaked in the diagonal elements (i.e., $D_{\alpha\beta}(x, y) \sim D_{\alpha\alpha}(x, x)\delta_{\alpha\beta}\delta_{x,y}$), which

r	δ_D (12×12)	δ_D (16×16)	δ_D (20×20)	index(D)
0.1	6.232×10^{-2}	1.792×10^{-2}	4.854×10^{-3}	1
0.2	4.818×10^{-3}	6.229×10^{-4}	7.750×10^{-5}	1
0.5	3.848×10^{-4}	2.434×10^{-5}	1.559×10^{-6}	1
0.8	1.698×10^{-4}	8.268×10^{-6}	4.976×10^{-7}	1
1.0	1.618×10^{-4}	6.596×10^{-6}	3.432×10^{-7}	1
1.2	3.448×10^{-4}	1.318×10^{-5}	5.308×10^{-7}	1
1.5	1.688×10^{-3}	1.243×10^{-4}	8.783×10^{-6}	1
2.0	9.940×10^{-3}	1.569×10^{-3}	2.348×10^{-4}	1
5.0	0.1599	9.314×10^{-2}	5.088×10^{-2}	1

Table 2: The deviation of the chiral anomaly function, δ_D [Eq. (24)], versus $R(x, y) = r \delta_{x,y}$, for lattice sizes, 12×12 , 16×16 and 20×20 respectively. The background gauge field has constant field tensors with topological charge $Q = 1$. The index of D is always equal to one in each case.

is unphysical since it does not respond properly to the background gauge field (e.g., the chiral anomaly is incorrect even though the index of D is equal to the background topological charge). In Table 2, we list the deviation of the chiral anomaly function, δ_D , versus $R(x, y) = r \delta_{x,y}$, on a 12×12 lattice (the second column), in a constant background gauge field with topological charge $Q = 1$. We see that the proper values of r are approximately in the range $0.5 \sim 1.2$, where D can reproduce the continuum chiral anomaly precisely. Next we investigate how the lattice size affects the range of proper values of r . The results of δ_D for lattice sizes 16×16 and 20×20 are listed in the third and the fourth columns in Table 2. They clearly show that the lower bound of r can be pushed to a smaller value, ~ 0.2 , when the size of the lattice is increased to 20×20 . Therefore, it suggests that the chiral limit ($r \rightarrow 0$ and $D \rightarrow D_c$) can be approached by decreasing the value of r while increasing the size of the lattice, at finite lattice spacing. This provides a *nonperturbative definition of the chiral limit* for any D of the general solution (3) with D_c satisfying the necessary physical requirements [3].

It is evident that the range of proper values of r also depends on the background gauge configuration. However, we suspect that when the background gauge configuration becomes very rough, there may not exist any values of r such that the chiral anomaly function is in good agreement with the Chern-Pontryagin density. We intend to return to this question in a later publication.

In summary, we have clarified the role of R in the general solution (3) of the Ginsparg-Wilson relation. It provides a topologically invariant transformation which transforms the chirally symmetric and nonlocal D_c into a local D which satisfies the GW relation, the exact chiral symmetry on the lattice. Having R local in the position space is a necessary condition to ensure the absence of ad-

ditive mass renormalization in the fermion propagator, as well as to produce a local D , which is vital for obtaining the correct chiral anomaly. Our numerical results strongly suggest that the optimal form of R is $R(x, y) = r \delta_{x, y}$. The range of proper values of r depends on the background gauge configuration as well as the size of the lattice, $L = Na$. In the limit $N \rightarrow \infty$, for smooth gauge backgrounds, the lower bound of proper values of r goes to zero, thus the chiral limit ($r \rightarrow 0$ and $D \rightarrow D_c$) can be approached nonperturbatively at finite lattice spacing.

Acknowledgement

I would like to thank all participants of Chiral '99, workshop on chiral gauge theories (Taipei, Sep. 13-18, 1999), for their stimulating questions and interesting discussions. I am also indebted to Herbert Neuberger for his helpful comments on the first version of this paper. This work was supported by the National Science Council, R.O.C. under the grant number NSC89-2112-M002-017.

References

- [1] P. Ginsparg and K. Wilson, Phys. Rev. D25 (1982) 2649.
- [2] T.W. Chiu and S.V. Zenkin, Phys. Rev. D59 (1999) 074501.
- [3] T.W. Chiu, Phys. Lett. B445 (1999) 371.
- [4] T.W. Chiu, "A construction of chiral fermion action", hep-lat/9908023, to be published in Phys. Lett. B.
- [5] H.B. Nielsen and N. Ninomiya, Nucl. Phys. B 185 (1981) 20 [E: B195 (1982) 541]; B193 (1981) 173.
- [6] T.W. Chiu, Nucl. Phys. B (Proc. Suppl.) 73 (1999) 688, hep-lat/9809085.
- [7] T.W. Chiu, "Fermion determinant and chiral anomaly on a finite lattice", hep-lat/9906007.
- [8] H. Neuberger, Phys. Lett. B417 (1998) 141; B427 (1998) 353.
- [9] T.W. Chiu, Phys. Rev. D58 (1998) 074511.

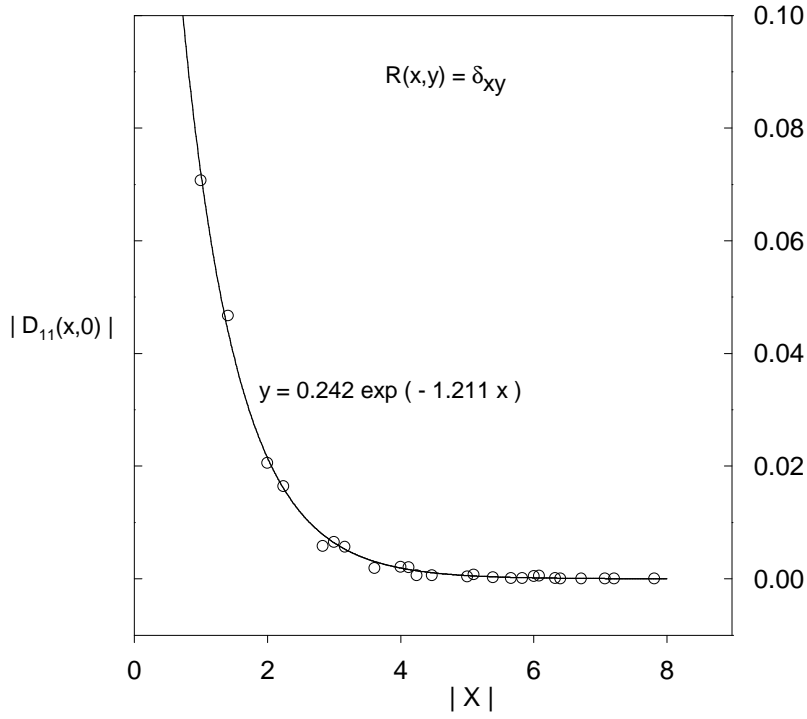


Figure 1: One of the Dirac components of $D(x,0)$, $|D_{11}(x,0)|$, is plotted as a function of $|x|$ for $R(x,y) = \delta_{xy}$. The lattice is 12×12 with periodic boundary conditions. The constant background gauge field has topological charge $Q = 1$. All data points at the same distance $|x|$ from the origin have been averaged. The solid line is an exponential fit to the data points. The same decay constant also fits very well for all other Dirac components of $D(x,y)$ and for any reference point y .

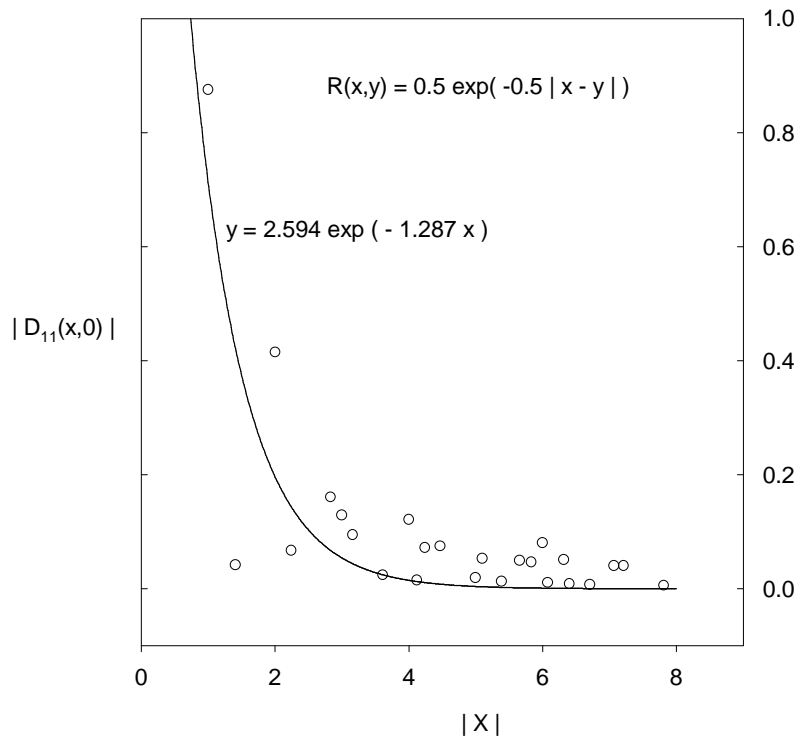


Figure 2: One of the Dirac components of $D(x,0)$ is plotted as a function of $|x|$ with $R(x,y) = 0.5 \exp(-0.5|x - y|)$. Other descriptions are the same as Fig. 1. The non-localness of D is shown clearly.

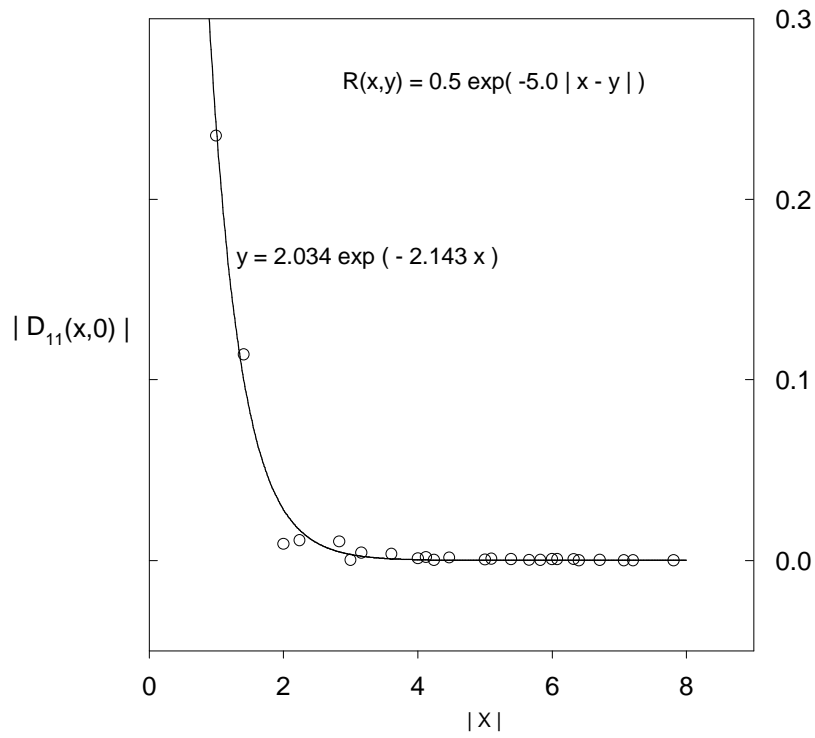


Figure 3: One of the Dirac components of $D(x,0)$ is plotted as a function of $|x|$ with $R(x,y) = 0.5 \exp(-5.0|x - y|)$. Other descriptions are the same as Fig. 1.

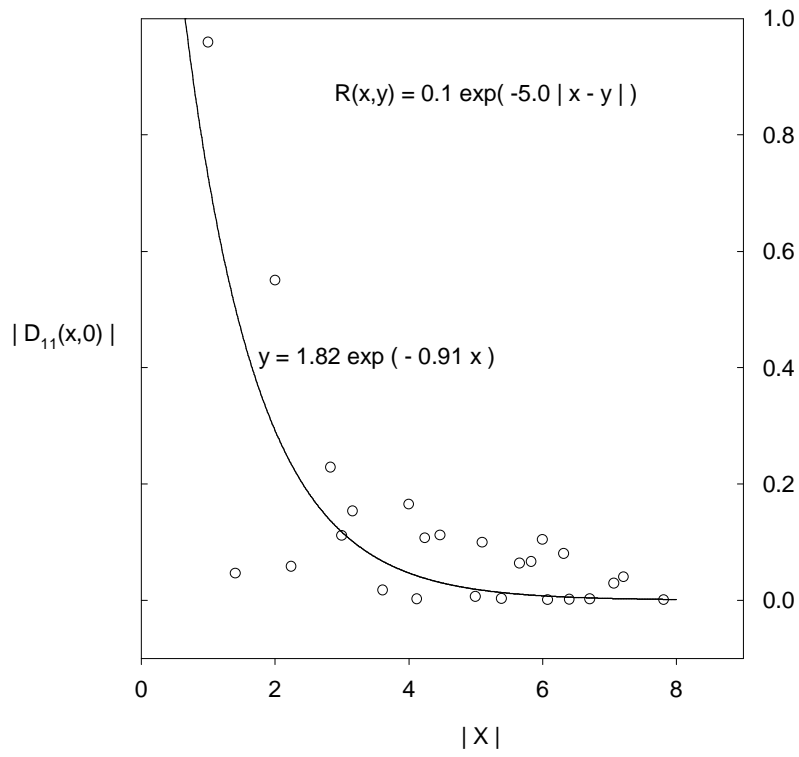


Figure 4: One of the Dirac components of $D(x,0)$ is plotted as a function of $|x|$ with $R(x,y) = 0.1 \exp(-5.0|x - y|)$. Other descriptions are the same as Fig. 1. The non-localness of D is shown clearly.